DOCUMENT RESUME

ED 268 178 TM 860 244

AUTHOR Cantor, Jeffrey A.; Hobson, Edward N.
TITLE The Development of a Model for Construction of Criterion Referenced System Achievement Tests for the Strategic Weapon System Training Program.
PUB DATE Apr 86
NOTE 17p.; Paper presented at the Annual Meeting of the American Educational Research Association (70th, San Francisco, CA, April 16-20, 1986).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Achievement Tests; Armed Forces; *Criterion Referenced Tests; Military Personnel; *Military Training; Models; Performance Factors; Postsecondary Education; Questionnaires; *Scores; *Test Construction; *Vocational Evaluation
IDENTIFIERS *Criterion Referenced System Achievement Tests; Ebel Method; Strategic Planning; *Strategic Weapons System Training Program

ABSTRACT

The development of a test design methodology used to construct a criterion-referenced System Achievement Test (CR-SAT) for selected Naval enlisted classification (NEC) in the Strategic Weapon System (SWS) of the United States Navy is described. Subject matter experts, training data analysts and educational specialists developed a comprehensive list of jobs/tasks and procedures necessary to the functioning of the NEC. This list resulted in a questionnaire of 360 items. Fleet technicians set performance standards, based on the Ebel method, for each job/task or procedure. Each group of test items were rated and summarized on a master matrix, and data findings were reduced to a percentage. Overall performance standards (cut scores) for each group of items were established and the test scoring algorithm was programmed into a computer. After testing, score reports are generated that detail technician performance in each of the major equipment categories. These reports describe performance in terms of the minimum acceptable cut scores, and in terms of overall fleet performance. This model provided a framework for a program evaluation which can supply feedback for management of organizational and personnel effectiveness, adjustments in training, and indications of new areas necessary for training. (PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED268178

The Development of a Model for Construction of Criterion Referenced System Achievement Tests for the Strategic Weapon System Training Program

Jeffrey A. Cantor
DDL OMNI Engineering
and
Edward N. Hobson
Military Training Associates

Presented at the annual meeting of the
American Educational Research Association
San Francisco, April, 1986

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
J. A. Cantor
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TM 860 244

Abstract

A model test design methodology for the U.S. Navy's Strategic Weapon System Training Program is described. The objective of this test design process is a criterion-referenced System Achievement Test. This process permits criterion inputs from current fleet personnel respective to performance standards. The reporting format highlights fleet performances in terms of individual technician scores, fleet performances, and indicates necessary areas for training.

The Development of a Model for Construction of Criterion Referenced System Achievement Tests for the Strategic Weapon System Training Program

Background

The U.S. Navy's Strategic Weapon System has a unique purpose: to prevent a nuclear attack on the United States. The effectiveness of the SWS as a deterrent to such an enemy attack and to global war is dependent upon well-trained officers and enlisted personnel. As a nuclear submarine on patrol is a self-contained entitity, the crew receives no outside assistance other than one-way information abd direction from higher command. Therefore, crew members must be well qualified, reliable, and self-sufficient in operating and maintaining the submarine and its weapon system. The goal of the SWS Training Program (SWSTP) is to provide comprehensive, cost-effective training in support of maintaining the SWS, its subsystems, and equipment, in a high state of opertional readiness.

Objective

The purpose of this paper is to describe the development of the test design methodology used to construct a criterion-referenced System Achievement Test for selected Naval enlisted classifications (NEC) in the Strategic Weapon System (SWS) of the U.S. Navy. The SWSTP goal since its inception has been to provide comprehensive, all encompassing individual and group training to ensure that the SWS can be operated and maintained in a high state of readiness. To meet this goal, the SWSTP has developed into an elaborate system of program objectives, training pipelines, curricula/courses and testing and evaluation procedures. An intregal part of this training program is the means by which it is kept current with the changing state-of-the-art hardware and operational requirements. One change to meet these needs was a movement towards a new methodology for the construction of a test design specification for the criterion-referenced System Achievement Test (C-R SAT).

Perspective

The C-R SAT is administered periodically to all SWS technicians to measure overall system knowledge and skill levels based on the specified jobs/tasks and procedures identified as essential to the performance of each NEC. The primary purpose of the C-R SAT is to evaluate the training systems; its secondary purpose is to assess the knowledge and skill of the personnel tested. Previous specifications for the SAT had been specified, generally, in terms of the format and content of retired versions of that SAT within a given NEC. This procedure did not allow for control in terms of overall coverage of all requisite jobs/tasks and procedures, nor did it allow for updating of information coverage necessary as a result of changes in technology of the equipment and procedures.

BEST COPY AVAILABLE

Methodology

The project undertaken to design a methodology capable of producing a test design specification procedure which would allow for inputs for adequate content coverage, performance standards, and changes in technology and also be self-sustaining in a dynamic NAVY personnel system was of enormus magnitude. The project team included contractor subject matter experts, training data analysts and educational specialists. The team also included SWS Fleet personnel.

The methodology devised for the new test design specification is based on the use of a committee of technical experts with current knowledge of the jobs/tasks and procedures, and the performance standards necessary for successful performance in the NEC for which the specification is to be constructed. The project leaders, in preparation for the committee activity, would prepare a comprehensive listing of those jobs/tasks and procedures identified as necessary to the function of the NEC. This list would be compiled from equipment documentation literature, procedural operations manuals and technician interviews (Figure 1).

Insert Figure 1 Here

This list is organized into a series of questionnaires ordered by equipment, capable of soliciting descriptive information about the performance of each of the jobs/tasks and procedures as perfromed by the individual technician in regard to the consequences of performace of the job/task or procedure; and the difficulty of performance of each job/task or procedure. A master listing of equipments is also prepared soliciting information from the technician about the performance of corrective maintenance procedures, routine maintenance procedures or casualty operating procedures. Information is requested from the technician respective to the necessity to know theory of operation and theory of interface description of the equipment as it relates to the ships systems. Information is then solicited at a second workshop regarding the acceptable standards of performance for each of the jobs/tasks and procedures.

Data Collection

The procedures used to collect data are as follows. The project management team request from the NAVY Special Projects Office (SSP) assistance in obtaining appropriate fleet personnel to serve as subject matter experts. These personnel will be representative of the NEC and will be distributed evenly between junior and supervisor level technicians, as follows.

Insert Figure 2 Here

BEST COPY AVAILABLE

A three to five day workshop is planned during which the committee will work individually and in groups (junior-supervisor) in order to provide technical input. A second workshop is planned and conducted to solicit performance standards input. Using a second group of technicians, the test design specification is developed and approved, and a prototype test is generated. The first process involves review of the nature of the equipment list according to performances in each of the procedures and descriptive areas. It is through this process that a hierarchial ranking of equipment will result. Accordingly if a technician was to indicate the performance of one or more jobs/tasks or procedures in each of the eight categories for an equipment, that equipment would receive eight points. The scoring algorethm involves totaling points for each equipment for each of the participating technicians.

Two separate accountings are maintained- one for junior and one for supervisor. Two separate test design specifications will be produced. Having completed the scoring, a hierarchial ranking is established. The equipment receiving the highest numerical score is ranked one, and so forth.

Next, the assignment of the number of test items per equipment is determined. The project team preagreed with SSP Management that a SAT would consist of 360 items. The dispersion of these items were into equipment areas, administration and security, casualty operating procedures and watchstanding requirements. The equipment section on the first SAT consisted of 22 equipments. A decision was made that only the top 15 equipments would be included on the test design specification, and that no equipment would receive less than 10 items. After the Committee completes the first round of equipment ratings, the project team tabulates the data, ranks the equipments and proportions the test items to each equipment based on hierarchial order.

At this point the Commitee begins to work through each of the equipments in hierarchial order as they were identified after the first phase of the workshop. The Committee is given directions and explanations of the rating scales. As the equipment packages are completed, the project team processes it and establishes hierarchial order of the categories and of the jobs/tasks and procedures within each category. Assignment of items for category is also completed at this point. After the equipment phases are completed, the Committee proceeds to group consensus seeking activity in order to review and rate the watchstanding qualifications and the casualty operating procedures. The last phase of the workshop is one that has proven to be of the most interest to the fleet committee members- the assignment of appropriate cognitive levels to the jobs/tasks and procedures.

The cognitive level assignment consists of the project leaders introducing the concept of Bloom's cognitive hierarchy to the fleet committee members and providing a sample exercise and critique session. At this point the commitee is given photocopies of the outcomes of the rated exercises (jobs/tasks and procedures). The membership embarks on individual assignment of cognitive levels culminating in their reconvening of two consensus-seeking groups to establish final cognitive levels.

BEST COPY AVAILABLE

The project leaders gather all of the data together and prepare a test design specification for each of the two levels of the NEC. The design specification will appear as an outline of the NEC in terms of jobs/tasks and procedures, numbers of test items per area and subarea and appropriate cognitive level for testing.

Insert Figure 3 Here

The second workshop is planned to solicit inputs from a new group of fleet technicians on performance standards (cut-scores) for each job/task or procedure. This workshop begins after the prototype test is constructed. The methodology used was adapted from the Ebel method for performance standard setting, which seeks inputs for test item respective to the minimally qualified technclans's ability to correctly respond to that item.

This process requires each technician to assign each test item to one cell of the Ebel Matrix - a matrix which provides for twelve categories; four levels of item importance and three levels of item difficulty. To assist in this exercise, a chart setting forth definitions of these item classifications is provided.

Insert Figures 4 & 5 Here

After each group of test items are rated, the workshop facilitator summarizes the groups' findings on a master matrix for the workshop. At this point the data findings are reduced to a percentage. The following chart illustrates the methodology used.

Insert Figure 6 Here

Once the overall cut-scores for each group of items have been established, the test scoring algorethm is programmed into the computer. A test version is now ready to be released to the fleet. The test cycle is two patrols in length. Therefore each crew of a boat will be exposed to each test version. The technicians will sit for the exam while on patrol at sea. The test packages will be returned for scoring as soon as possible after the testing is completed. Score reports are immediately generated, detailing

BEST COPY AVAILABLE

technician performances in each of the major equipment categories. These scoreing reports will describe a technciian's performances in terms of the minimum acceptable cut-scores, and in terms of the overall fleet performances. The following is an example of a typical score report format.

Insert Figure 7 Here

Conclusions

The completed C-R SAT process will provide a basis for fleet personnnel who are responsible for test administration to select appropriate test items in proper proportions, and to construct C-R tests as needed. The model also has the capability to identify other areas in the SWSTP needing attention. This model should serve to provide a framework for other large system training programs to begin to move to an program of evaluation which can serve to provide meaningful feedback for management of organizational and personnel effectiveness, adjustments in training and indications of new areas necessary for training. This program also provides one totally encompassing evaluation model which includes a test design process, a performance standard setting process and a training evaluation and feedback loop to the training planning design and development stages. Very few, if any, ongoing successful models of this kind exist, either within the military or civilian sectors.

BEST COPY AVAILABLE

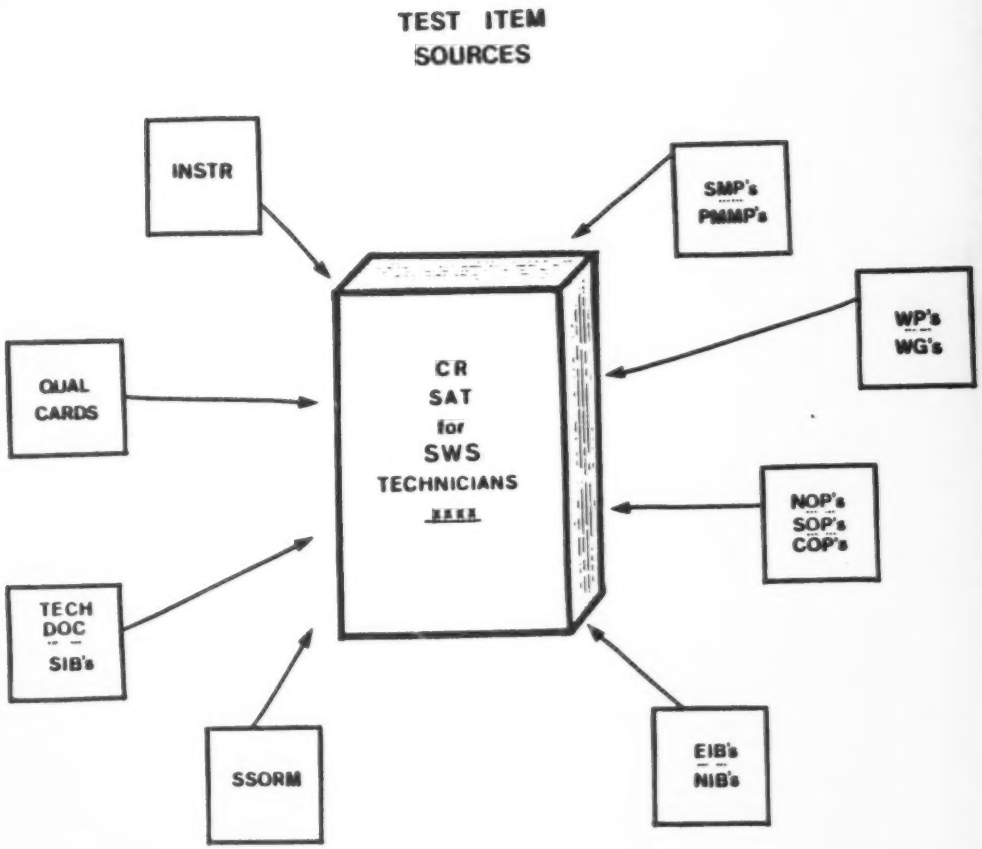
TEST ITEM
SOURCES
INSTR
SMP's
PMMP's
WP's
WG's
QUAL
CARDS
CR
SAT
for
SWS
TECHNICIANS
NOP's
SOP's
COP's
TECH
DOC
SIB's
SSORM
EIB's
NIB's

Figure 1.

WATCHSTANDER/JUNIOR TECHNICIAN

- E4 OR E5
- SUBMARINE QUALIFIED
- QUALIFIED ON JUNIOR WATCH
- COMPLETED ≥ 2 PATROLS
- COMPLETED ≥ 2 ADVANCED TRAINING COURSES

SUPERVISOR/SENIOR TECHNICIAN

- E5 OR ABOVE
- SUBMARINE QUALIFIED
- QUALIFIED AS A SUPERVISOR
- COMPLETED ≥ 3 PATROLS
- COMPLETED ≥ 3 ADVANCED TRAINING COURSES

Figure 2.

TRIDENT-I CNC TECHNICIAN

JUNIOR LEVEL SAT CONTENT

TEST AREA	Total Test Items	R	C	A	A/E
I. ADMINSTRATION/SECURITY	24	12	5	7	-
A. Security Requirements	2	2	-	-	-
B. Administrative Requirements	22	10	5	7	-
II. EQUIPMENT	135	77	58	-	-
A. CNC	40	27	13	-	-
1. Functional Operation	8	6	2	-	-
2. Interface Description	6	2	4	-	-
3. Corrective Maintenance	8	4	4	-	-
4. Standard Maintenance Procedures	8	5	3	-	-
5. Standard Operating Procedures	10	10	-	-	-
B. MTU	30	14	16	-	-
1. Functional Operation	6	4	2	-	-
2. Interface Description	5	2	3	-	-
3. Corrective Maintenance	6	3	3	-	-
4. Standard Maintenance Procedures	6	3	3	-	-
5. Standard Operating Procedures	7	2	5	-	-

Figure 3.

BEST COPY AVAILABLE

ITEM CLASSIFICATION DEFINITIONS

WHEN ASSESSING THE TEST ITEM, THINK IN TERMS OF THE MINIMALLY COMPETENT TECHNICIAN.

DIFFICULTY

- o EASY - THE TECHNICIAN WOULD IMMEDIATELY RECALL THE CORRECT ANSWER
- o MODERATE - THE TECHNICIAN WOULD NEED TO REVIEW THE ITEM ALTERNATIVES BEFORE IDENTIFYING THE CORRECT ANSWER.
- o HARD - THE TECHNICIAN MUST WRESTLE WITH THE ITEM ALTERNATIVES IN ORDER TO ELIMINATE INCORRECT RESPONSES.

RELEVANCE

- o ESSENTIAL - REPRESENTS KNOWLEDGE THAT IS OF UTMOST IMPORTANCE TO SUCCESSFULLY COMPLETING THE SUBMARINE'S MISSION.
- o IMPORTANT - REPRESENTS KNOWLEDGE THE TECHNICIAN NEEDS TO KNOW IN ORDER TO PROPERLY PERFORM HIS DUTIES.
- o ACCEPTABLE - THIS QUESTION REPRESENTS USEFUL/NICE-TO-KNOW INFORMATION.
- o QUESTIONABLE - THIS INFORMATION IS NOT REQUIRED IN ORDER TO DO THE JOB.

Figure 4.

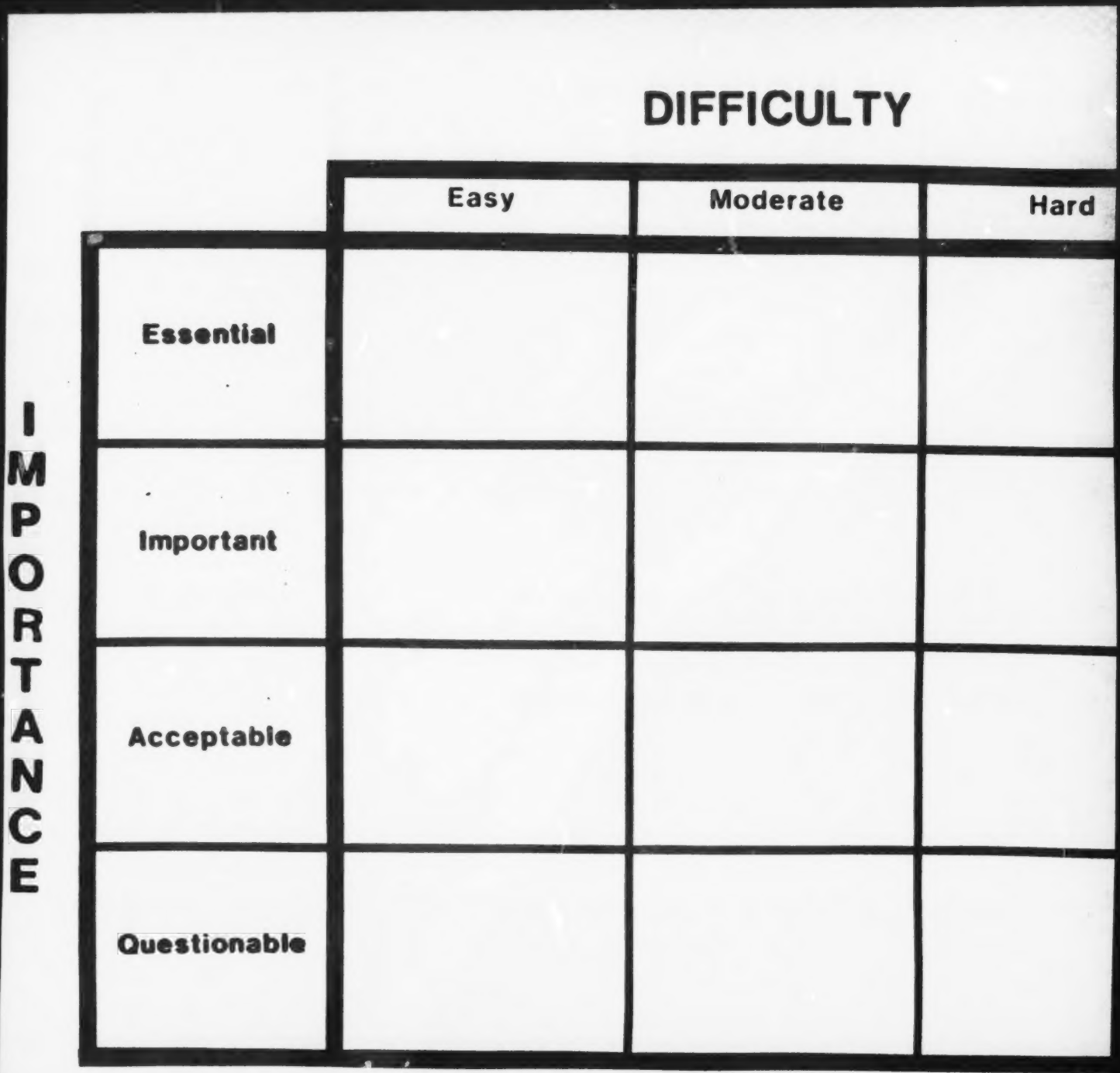
DIFFICULTY
Easy
Moderate
Hard
IMPORTANCE
Essential
Important
Acceptable
Questionable

Figure 5.

IMPORTANCE	DIFFICULTY: EASY	DIFFICULTY: MEDIUM	DIFFICULTY: HARD
ESSENTIAL	# OF ITEMS 10 PROPORTION .90	# OF ITEMS 15 PROPORTION .85	# OF ITEMS 8 PROPORTION .70
VERY IMPORTANT	# OF ITEMS 17 PROPORTION .80	# OF ITEMS 6 PROPORTION .75	# OF ITEMS 3 PROPORTION .60
NOT VERY IMPORTANT	# OF ITEMS 4 PROPORTION .65	# OF ITEMS 8 PROPORTION .50	# OF ITEMS 9 PROPORTION .50

CUT-OFF SCORE = (10 x .90) + (15 x .85) + (8 x .70) +
(17 x .80) + (6 x .75) + (3 x .60) +
(4 x .65) + (8 x .50) + (9 x .50) = 58.4

MAXIMUM POSSIBLE SCORE = 80.0

Figure 6.

07/10/95 SYSTEM ACHIEVEMENT TEST SSBN
NS3W/9501/03 Tested 07/09/95
SINS TECHNICIAN (NEC 3324) -- WATCHSTANDER Page 3

TEST SCORE SUMMARY

Rating/Rate... Name..... SSAN.....	GROUP AUG. SCORE	ET2	ET3
TOTAL TEST PERFORMANCE	72	70	75
A. ADMINISTRATION/ SECURITY	<69>	<64>	72
B. WATCH QUALIFICATIONS	75	71	79
C. EQUIPMENT	72	70	75
1. SINS	76	78	74
2. MSR	78	70	85
3. MARDAN	73	75	70
4. NOCC	74	<57>	89
5. Tape Reader	72	70	73
6. IBM Typewriter	<64>	70	<59>
7. Optical	<49>	<45>	<51>
8. MARDAN Test Set	75	80	70
9. NSSB	76	<62>	89

NOTE: A score of 70 or above indicates the examinee met or exceeded the minimum requirements for the total test, test area or subarea.

<> Score fell below the minimum requirement.

Figure 7.

BEST COPY AVAILABLE